Network Performance Testing

General-purpose network performance tests, such as latency and bandwidth benchmarks, do not represent the in vivo network usage of any particular application—thus, such tests may yield misleading results. Although special-purpose performance tests can provide more insight into application behavior they are often tedious to write and difficult to explain to others.

coNCePTuaL TO THE RESCUE

The Laboratory's Performance and Architecture Lab has developed a new tool that makes it easy to develop special-purpose network performance tests. Known as coNCePTuaL, this tool consists of a domain-specific programming language expressly developed for writing network benchmarks.

A REAL-WORLD EXAMPLE

We recently plotted a bandwidth curve and observed the characteristic "S" shape. However, we noted a performance discrepancy between the benchmarked bandwidth and that observed by another program. We believed the source of the discrepancy to be the message-buffer alignment. To test this hypothesis, we wrote the brief coNCePTuaL program shown in the next column.

coNCePTuaL enables a programmer to focus on the important parts of a network benchmark: communication and data acquisition. We used coNCePTuaL's "C + MPI" compiler backend to compile the program shown above and observed the performance shown below.

CONCLUSIONS

Network performance is extremely sensitive to message-buffer alignment. Page-aligned buffers yield suboptimal performance on both platforms. The InfiniBand platform is insensitive to receive-buffer alignment but has performance spikes on 128-byte offsets whereas the QsNet platform rather surprisingly sees the best performance when the send buffer is exactly 16 bytes ahead of the receive buffer.

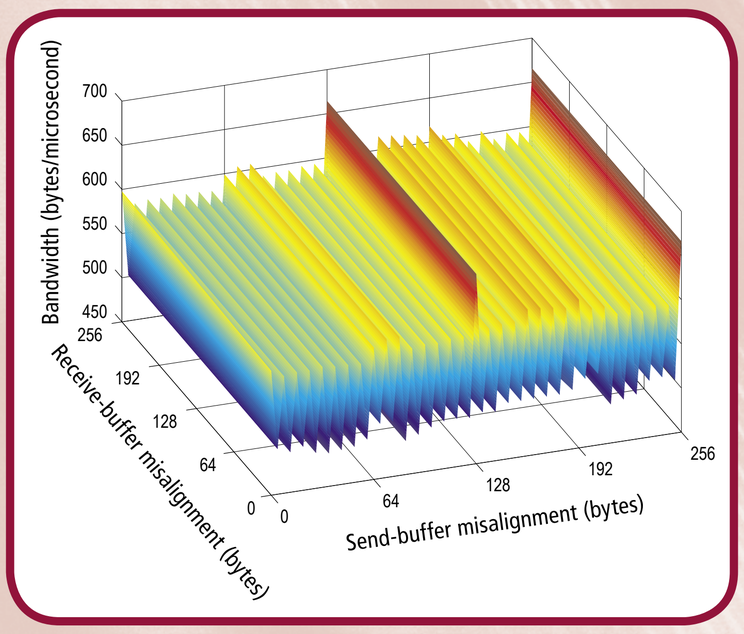
coNCePTuaL: MAKING IT EASY TO DEVELOP SPECIAL-PURPOSE NETWORK PERFORMANCE TESTS

coNCePTuaL HOME PAGE

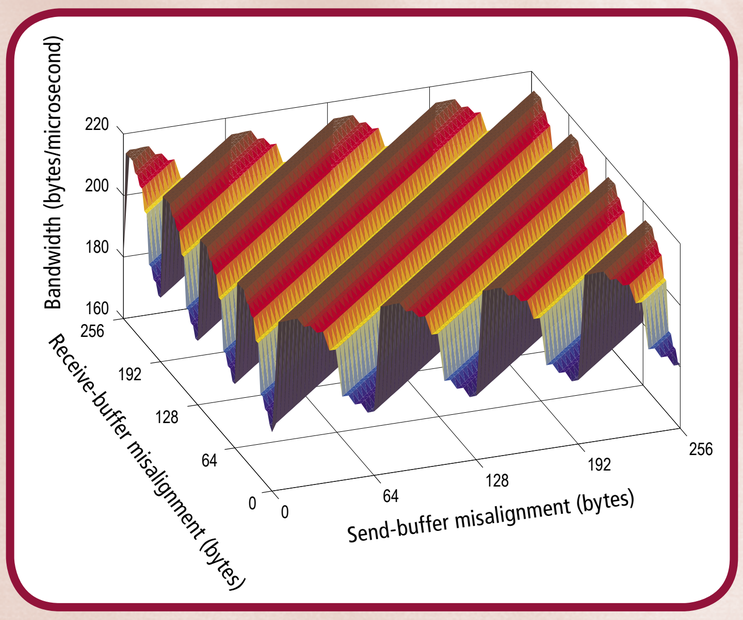
http://conceptual.sf.net/ ■

```
###############################################################################
# coNCePTuaL program to test all possible combinations of sender and receiver misalignment relative to a page boundary. #
###############################################################################

# Provide forward and backward compatibility as the coNCePTuaL language evolves.
Require language version "0.5.2".

# Parse the command line.
reps is "Number of repetitions" and comes from "--reps" or "-r" with default 100.
maxoffset is "Maximum offset in bytes from a page boundary" and comes from "--maxofs" or "-o" with default 16K.
msgsize is "Message size in bytes" and comes from "--size" or "-s" with default 1M.

# Ensure that we have a peer with whom to communicate.
Assert that "the latency test requires at least two tasks" with num_tasks>=2.

# Ping-pong using every combination of sender and receiver misalignments.
For each sendalign in {0, 4, 8, ..., maxoffset}
    for each recvalign in {0, 4, 8, ..., maxoffset} {
        task 0 outputs "Sender is offset by " and sendalign and "; receiver is offset by " and recvalign then
        all tasks synchronize then
        task 0 resets its counters then
        for reps repetitions plus 1 warmup repetition {
            task 0 sends a msgsize byte sendalign byte misaligned message to task 1 who receives it as a recvalign byte misaligned message then
            task 1 sends a msgsize byte sendalign byte misaligned message to task 0 who receives it as a recvalign byte misaligned message
        }
        then task 0 logs sendalign as "Sender misalignment (B)" and
                                 recvalign as "Receiver misalignment (B)" and
                                 total_bytes/elapsed_usecs as "Bandwidth (B/us)"
    }
```

Complete coNCePTuaL program for measuring bandwidth as a function of send- and receive-buffer alignment.



Mellanox InfiniBand 4X across a pair of dual 2.2 GHz Xeon nodes (PCI-X)



Quadrics QsNet across a pair of dual 1.3 GHz Itanium II nodes (PCI bridged to PCI-X)

PAL Performance and Architecture Lab
http://www.c3.lanl.gov/par_arch/

The World's Greatest Science Protecting America

Los Alamos NATIONAL LABORATORY EST.1943

LA-UR-04-7257